

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-298556

(43)Date of publication of application : 24.10.2000

(51)Int.Cl.

G06F 3/06

(21)Application number : 11-105030

(71)Applicant : NEC CORP

(22)Date of filing : 13.04.1999

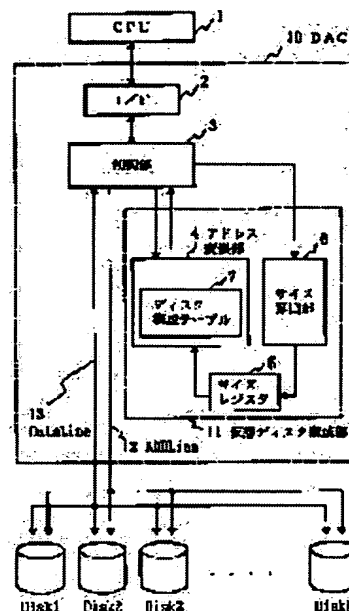
(72)Inventor : NAKAMURA NAOTAKA

(54) DISK ARRAY DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To construct a disk array having the same or almost the same capacity as that of all disk devices by using plural disk devices having respectively different capacity values.

SOLUTION: A size calculation part 6 in a disk array controller 10 determines the greatest common measure of capacity values of plural real disk devices Disk1 to DiskN as the size of a virtual disk device. A control part 3 sets up individual divided areas obtained by dividing the disk areas of respective real disk devices Disk1 to DiskN by the size as virtual disk devices and constructs a disk array by using plural virtual disk devices including plural virtual disk devices generated on the same disk device. When read and write requests are outputted from a master device to the disk array, an address conversion part 4 refers to a disk constitution table and converts the addresses of the virtual disk devices into the addresses of the real disk devices.



LEGAL STATUS

[Date of request for examination] 24.03.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3454183

[Date of registration] 25.07.2003

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(2)

特開2000-298556

1

【特許請求の範囲】

【請求項1】 少なくとも何れか2つの実ディスク装置の容量が異なっている複数の実ディスク装置を同じ容量のディスク領域に分割したときの個々のディスク領域を仮想ディスク装置とし、同じ実ディスク装置上に生成された複数の仮想ディスク装置を含め、同容量の複数の仮想ディスク装置を使ってディスクアレイを構成したディスクアレイ装置。

【請求項2】 少なくとも何れか2つの実ディスク装置の容量が異なっている複数の実ディスク装置と、該複数の実ディスク装置に接続された制御装置とから構成されたディスクアレイ装置において、前記制御装置は、

前記複数の実ディスク装置のディスク領域を分割するサイズを保持するサイズレジスタと、
前記複数の実ディスク装置を前記サイズで分割した個々のディスク領域を仮想ディスク装置とし、各仮想ディスク装置と前記実ディスク装置との対応関係を示すディスク構成テーブルを生成する手段と、

同じ実ディスク装置上に生成された複数の仮想ディスク装置を含め、複数の仮想ディスク装置を使ってディスクアレイを構成する手段と、

上位装置からのディスクアレイに対するリードおよびライト要求時、前記ディスク構成テーブルを参照して仮想ディスク装置に対するアドレスから対応する前記実ディスク装置に対するアドレスに変換する手段とを備えることを特徴とするディスクアレイ装置。

【請求項3】 生成された全ての仮想ディスク装置を使ってRAID0に相当するディスクアレイを構成した請求項2記載のディスクアレイ装置。

【請求項4】 マスタとなる仮想ディスク装置とスレーブとなる仮想ディスク装置とが同じ実ディスク装置上に存在しない条件を満足する範囲内で、できるだけ多くの仮想ディスク装置を使ってRAID1に相当するディスクアレイを構成した請求項2記載のディスクアレイ装置。

【請求項5】 生成された仮想ディスク装置の数が最小の実ディスク装置における仮想ディスク装置をパリティ用ディスク、他の実ディスク装置上に生成された仮想ディスク装置をデータ用ディスクにそれぞれ使ってRAID3に相当するディスクアレイを構成した請求項2記載のディスクアレイ装置。

【請求項6】 前記複数の実ディスク装置の容量の最大公約数を求める手段を備え、該手段で求められた値を前記サイズレジスタに設定する請求項2、3、4または5記載のディスクアレイ装置。

【請求項7】 最小サイズを保持するサイズ設定レジスタと、前記複数の実ディスク装置の容量の最大公約数を求める手段とを備え、該手段で求められた値が前記最小サイズより小さくなる場合には前記最小サイズを前記サ

2

イズレジスタに設定し、それ以外は前記求められた値を前記サイズレジスタに設定する請求項2、3、4または5記載のディスクアレイ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明はディスクアレイ装置に関し、特に容量の異なるディスク装置を使ってディスクアレイを構築する技術に関する。

【0002】

【従来の技術】 複数のディスク装置にデータを分散させて記憶させるディスクアレイ装置は、データをどのように分散して記憶させるかによって、RAID0～RAID5などの構成に分類されるが、何れも、全て同じ容量のディスク装置を使って構築されるのが一般的であった。異なった容量のディスク装置を使用してディスクアレイを構成することも可能であるが、その場合には、各ディスク装置の容量がいくらであっても、ディスクアレイを構成するうちの最小容量のディスク装置と同じ容量としてしか使用されず、ディスク容量の使用効率が著しく低下していた。

【0003】 このような事情に鑑み、異なる容量のディスク装置を使ってディスクアレイを構築する際に、従来のディスクアレイとして使用できなかったディスク領域もディスクアレイに使用できるようにする技術として、特開平8-249132号公報に記載される技術がある。この技術（以下、従来技術と称す）では、各ディスク装置におけるディスク領域の先頭から最小容量のディスク装置と同じ容量分のディスク領域までを使って1つのディスクアレイを構成し、最小容量以外のディスク装置における余ったディスク領域を別のディスクアレイに利用する。例えば、1Gバイトのディスク装置4台と、500Mバイトのディスク装置1台とが存在する場合、1Gバイトのディスク装置の先頭500Mバイトのディスク領域と500Mバイトのディスク装置の全ディスク領域とを使って、500M×5=2.5Gバイトのディスクアレイを構成し、1Gバイトのディスク装置の残り500Mバイトのディスク領域を使って、500M×4=2Gバイトの別のディスクアレイを構成する。同様な技術は特開平9-120342号公報にも記載されている。

【0004】

【発明が解決しようとする課題】 上述した従来技術によれば、ディスクアレイとして従来使用できなかったディスク領域もディスクアレイに使用することができ、ディスク領域の使用効率を向上させることができる。しかしながら、構築可能な1つのディスクアレイの最大容量は、「最小容量のディスク容量×ディスク装置台数」に制限されてしまい、ディスクアレイの最大の特徴である大容量化が実現できない。

【0005】 そこで本発明の目的は、異なった容量のディスク装置を使って、全ディスク装置の容量の合計と同

(3)

特開2000-298556

3

じか、それにはほぼ近い容量のディスクアレイを構築できるようにすることにある。

【0006】

【課題を解決するための手段】本発明は上記の目的を達成するために、少なくとも何れか2つの実ディスク装置の容量が異なっている複数の実ディスク装置を同じ容量のディスク領域に分割したときの個々のディスク領域を仮想ディスク装置とし、同じ実ディスク装置上に生成された複数の仮想ディスク装置を含め、同容量の複数の仮想ディスク装置を使ってディスクアレイを構成する。より具体的には、少なくとも何れか2つの実ディスク装置の容量が異なっている複数の実ディスク装置と、該複数の実ディスク装置に接続された制御装置とから構成されたディスクアレイ装置において、前記制御装置は、前記複数の実ディスク装置のディスク領域を分割するサイズを保持するサイズレジスタと、前記複数の実ディスク装置を前記サイズで分割した個々のディスク領域を仮想ディスク装置とし、各仮想ディスク装置と前記実ディスク装置との対応関係を示すディスク構成テーブルを生成する手段と、同じ実ディスク装置上に生成された複数の仮想ディスク装置を含め、複数の仮想ディスク装置を使ってディスクアレイを構成する手段と、上位装置からのディスクアレイに対するリードおよびライト要求時、前記ディスク構成テーブルを参照して仮想ディスク装置に対するアドレスから対応する前記実ディスク装置に対するアドレスに変換する手段とを備えている。

【0007】また本発明は、生成された全ての仮想ディスク装置を使ってRAID0に相当するディスクアレイを構成する。

【0008】また本発明は、マスタとなる仮想ディスク装置とスレーブとなる仮想ディスク装置とが同じ実ディスク装置上に存在しない条件を満たす範囲内で、できるだけ多くの仮想ディスク装置を使ってRAID1に相当するディスクアレイを構成する。

【0009】また本発明は、生成された仮想ディスク装置の数が最小の実ディスク装置における仮想ディスク装置をパリティ用ディスク、他の実ディスク装置上に生成された仮想ディスク装置をデータ用ディスクにそれぞれ使ってRAID3に相当するディスクアレイを構成する。

【0010】また本発明は、更に、前記複数の実ディスク装置の容量の最大公約数を求める手段を備え、該手段で求められた値を前記サイズレジスタに設定する。

【0011】また本発明は、更に、最小サイズを保持するサイズ設定レジスタと、前記複数の実ディスク装置の容量の最大公約数を求める手段とを備え、該手段で求められた値が前記最小サイズより小さくなる場合には前記最小サイズを前記サイズレジスタに設定し、それ以外は前記求められた値を前記サイズレジスタに設定する。

【0012】本発明のディスクアレイ装置にあっては、

4

複数の実ディスク装置を同じ容量のディスク領域に分割したときの個々のディスク領域を仮想ディスク装置とし、同じ実ディスク装置上に生成された複数の仮想ディスク装置を含め、同容量の複数の仮想ディスク装置を使ってディスクアレイを構成するため、全ディスク装置の容量の合計と同じか、それにはほぼ近い容量のディスクアレイを構築することが可能となる。

【0013】

【発明の実施の形態】次に本発明の実施の形態の例について図面を参照して詳細に説明する。

【0014】図1は本発明の一実施の形態の原理説明図である。今、図1(a)に示されるように、容量の異なる2つの実ディスク装置D1、D2を使ってディスクアレイを構築する場合を考える。ここで、実ディスク装置D1の容量は実ディスク装置D2の1.5倍あるものとする。従来は、容量の小さい実ディスク装置D2の全ディスク領域と、実ディスク装置D1の先頭から実ディスク装置D2の容量分のディスク領域とで、1つのディスクアレイを構成しており、実ディスク装置D1の残りのディスク領域は使用されていなかった。特開平8-249132号公報では、この使用されていなかった実ディスク装置D1の残りのディスク領域を別のディスクアレイとして活用する。従って、ディスク領域の使用効率は向上するが、実ディスク装置D2の容量の2倍のディスクアレイと、実ディスク装置D2の容量の半分のディスクアレイとの2つのディスクアレイができていた。

【0015】本発明の場合は、図1(b)に示すように、実ディスク装置D1、D2のディスク領域を、例えば各容量の最大公約数のサイズで分割し、その各々の分割領域を仮想ディスク装置D11～D13、D21、D22とし、これらの仮想ディスク装置D11～D13、D21、D22の各々を1つの実ディスク装置に見立てて、1つのディスクアレイを構成する。この結果、ディスクアレイの容量は実ディスク装置D1、D2の容量の合計値に等しくなり、大容量のディスクアレイが構築できる。以下、本発明の実施の形態の実施例について説明する。

【0016】

【実施例】図2を参照すると、本発明のディスクアレイ装置の一実施例は、上位装置であるCPU1に接続されたディスクアレイコントローラ(DAC)10と、このディスクアレイコントローラ10にアドレス線12及びデータ線13を通じて接続された複数の実ディスク装置Disk1～DiskNとから構成されている。ここで、複数の実ディスク装置Disk1～DiskNのうち、少なくとも何れか2つの実ディスク装置は互いに容量が異なっている。

【0017】また、ディスクアレイコントローラ10は、制御部3と、この制御部3とCPU1との間に接続されたインターフェース2と、制御部3に接続された仮

(4)

特開2000-298556

5

6

想ディスク構成部11とを含んでいる。

【0018】更に、仮想ディスク構成部11は、制御部3に接続されたアドレス変換部4およびサイズ算出部6と、アドレス変換部4とサイズ算出部6との間に設けられたサイズレジスタ5とを有し、アドレス変換部4はディスク構成テーブル7を内蔵している。

【0019】図3を参照すると、制御部3の一実施例は、インターフェース2に接続されたMPU301と、このMPU301が実行すべきプログラムや各種の情報を保持するRAM303と、MPU301に接続されたキャッシュ302、セクタ304、インターフェース306、307と、セクタ304と各実ディスク装置Disk1～DiskNとの間に設けられたディスクインターフェース305-1～305-Nとを有し、アドレス変換部4とはインターフェース306を介して、またサイズ算出部6とはインターフェース307を介して、それぞれ接続されている。

【0020】図4は複数の実ディスク装置を使ってディスクアレイを構築する手順の一例を示すフローチャートである。まず、ディスクアレイに使う実ディスク装置を決定する(ステップS1)。これは、システム管理者が決定して、CPU1経由でディスクアレイコントローラ10の制御部3におけるMPU301に通知することで行う。また、ディスクアレイコントローラ10に実際に接続されている実ディスク装置すべてをディスクアレイに使う場合には、MPU301がセクタ304、ディスクインターフェース305-1～305-Nを介して実装されている実ディスク装置を認識することで、自動的に決定することも可能である。以下では、実ディスク装置Disk1～DiskNの全てを使って1つのディスクアレイを構築するものとする。

【0021】次に、ディスクアレイに使う各実ディスク装置の番号(ID)と容量を取得し、RAM303に一次的に保存する(ステップS2)。これは、制御部3のMPU301が各実ディスク装置Disk1～DiskNの番号と容量を取得してRAM303に記憶しても良く、システム管理者自身がCPU1経由でディスクアレイコントローラ10の制御部3内のRAM303に記憶させるようにしても良い。

【0022】次に、仮想ディスク装置の容量(サイズ)を決定する(ステップS3)。これは、ディスクアレイに使う各実ディスク装置の容量をRAM303から読み出して、制御部3からインターフェース307を通じてサイズ算出部6に与えることで、サイズ算出部6で決定させる。サイズ算出部6は、本実施例の場合、与えられた各実ディスク装置の容量の最大公約数を求め、その値を仮想ディスク装置のサイズに決定し、MPU301に通知すると共にサイズレジスタ5に記憶させる。

【0023】次に、制御部3のMPU301は、通知されたサイズで各実ディスク装置Disk1～DiskN

のディスク領域を分割したときの個々の分割領域を1つの仮想ディスク装置とし、仮想ディスク装置と実ディスク装置との対応関係を示すディスク構成テーブルを作成する(ステップS4)。図5にディスク構成テーブルの例を示す。ディスク構成テーブルでは、各仮想ディスク装置毎に、それがどの実ディスク装置上に生成されているかを示す実ディスク装置の番号と、その実ディスク装置の何番目の仮想ディスク装置であるかを示す通し番号とが記録されている。通し番号は、実ディスク装置のディスク領域の先頭のものから順に1、2、…と採番される。例えば、図6に示すように、或る実ディスク装置の先頭から(N-1)×サイズよりN×サイズまでのディスク領域に対応する仮想ディスク装置D1Nの場合、通し番号はNである。以上のようにして作成されたディスク構成テーブルは、制御部3のRAM303に記憶されると共に、インターフェース306を介してアドレス変換部4に送られ、ディスク構成テーブル7としてアドレス変換部4に保持される。

【0024】次に、以上のようにして生成された複数の仮想ディスク装置を使って1つのディスクアレイを構築すべく、ディスクアレイ構成情報を作成し、RAM303に記憶する(ステップS5)。この処理は、実ディスク装置によってディスクアレイを構築する際とはほぼ同じ処理であり、ただ、実ディスク装置の代わりに仮想ディスク装置を使用する点異なる。構築できるディスクアレイは、本実施例の場合、RAID0、RAID1、RAID3の3種類である。何れを作成するかは、例えばCPU1を通じてシステム管理者がMPU301に指定する。勿論、作成するRAIDの種類がRAM303に記憶された制御パラメータで事前に指定されている場合には、システム管理者からの指定は不要である。以下、各RAID毎に具体的なディスクアレイの構成例を説明する。

【0025】なお、説明の便宜上、ディスクアレイに使う実ディスク装置はD1、D2、D3の3つであり、容量は実ディスク装置D3が最も小さく、実ディスク装置D2は実ディスク装置D3の倍の容量があり、実ディスク装置D1は実ディスク装置D2のさらに倍の容量があるものとする。また、サイズ算出部6では、最大公約数として実ディスク装置D3の容量と同じ値を決定し、そのため、実ディスク装置D1は4分割されて4つの仮想ディスク装置D11、D12、D13、D14が生成され、実ディスク装置D2は2分割されて2つの仮想ディスク装置D21、D22が生成され、実ディスク装置D3は分割されずに1つの仮想ディスク装置D31が生成されたものとする。

【0026】(1) RAID0

RAID0の場合、図7に示すように、実ディスク装置D1上の4つの仮想ディスク装置D11～D14と、実ディスク装置D2上の2つの仮想ディスク装置D21、

(5)

特開2000-298556

7

8

D22と、実ディスク装置D3上の1つの仮想ディスク装置D31との合計7つの仮想ディスク装置すべてを使ってディスクアレイを構築する。そして、各仮想ディスク装置のディスク領域に対し図7の1からnまでに示されるような論理アドレス付けが行われ、ディスクアレイ構成情報として、例えば図8に示すように、各論理アドレス毎に、その論理アドレスが指し示す領域が存在する仮想ディスク装置の番号と、ディスク内アドレス（先頭からのオフセットアドレス）とを記録した情報が生成される。なお、Sは仮想ディスク装置のサイズ、Lは1論理アドレス当たりのアドレス増分値である。

【0027】(2) RAID1

RAID1の場合、マスタとなる仮想ディスク装置とスレーブとなる仮想ディスク装置とをペアにするので、偶数個の仮想ディスク装置を使用する。また、マスタとなる仮想ディスク装置とスレーブとなる仮想ディスク装置とは同じ実ディスク装置上に存在しないようにする。これらの条件を満足する範囲内で、できるだけ多くの仮想ディスク装置を使ってRAID1に相当するディスクアレイを構築する。今の例では、図9に示すように、仮想ディスク装置D14以外の6つの仮想ディスク装置D11~D13、D21、D22、D31を使ってディスクアレイを構築し、D11とD21、D12とD31、D13とD22をそれぞれペアとする。そして、各仮想ディスク装置D11~D13、D21、D22、D31のディスク領域に対し図9の1からnまでに示されるような論理アドレス付けが行われ、ディスクアレイ構成情報として、例えば図10に示すように、各論理アドレス毎に、マスタとなる仮想ディスク装置の番号と、スレーブとなる仮想ディスク装置の番号と、ディスク内アドレス（先頭からのオフセットアドレス）とを記録した情報が生成される。

【0028】なお、生成された仮想ディスク装置が奇数個であった場合に、どの1つの仮想ディスク装置を未使用とするかは、CPU1を通じてシステム管理者がMPU301に指定しても良く、また、MPU301自身が、例えば最大数の仮想ディスク装置が生成された実ディスク装置の最後の仮想ディスク装置（ディスク領域の先頭から最も遠くにあるもの）を未使用にするようにしても良い。さらに、ペアにする仮想ディスク装置の指定はCPU1を通じてシステム管理者がMPU301に指定しても良く、MPU301自身が、マスタとなる仮想ディスク装置とスレーブとなる仮想ディスク装置とが同じ実ディスク装置上に存在しない条件を満足する範囲内で、できるだけ多くのペアを求める組み合わせ問題を解くことで決定しても良い。

【0029】(3) RAID3

RAID3の場合、MPU301は、生成された仮想ディスク装置の数が最小の実ディスク装置における仮想ディスク装置をパリティ用ディスク、他の実ディスク装置

上に生成された仮想ディスク装置をデータ用ディスクにそれぞれ使ってRAID3に相当するディスクアレイを構築する。従って、今の例では、図11に示すように、実ディスク装置D3上の仮想ディスク装置D31がパリティ用ディスク、実ディスク装置D3以外の実ディスク装置D1、D2上の仮想ディスク装置D11~D14、D21、D22がデータ用ディスクとして使ったディスクアレイが構築される。勿論、CPU1を通じてシステム管理者がパリティ用ディスクとして使用する仮想ディスク装置、データ用ディスクとして使用する仮想ディスク装置をMPU301に指定するようにしても良い。そして、各仮想ディスク装置D11~D14、D21、D22、D31のディスク領域に対し図11のA-1からn-6までに示されるような論理アドレス付けが行われ、ディスクアレイ構成情報として、例えば図12に示すように、各論理アドレス毎に、データ用仮想ディスク装置の番号およびそのディスク内アドレス（先頭からのオフセットアドレス）と、パリティ用仮想ディスク装置の番号およびそのディスク内アドレス（先頭からのオフセット）とを記録した情報が生成される。

【0030】次に、上述のようにして構築されたディスクアレイに対するリード、ライト時の動作を説明する。

【0031】(1) RAID0

CPU1から論理アドレスを指定したリード要求またはライト要求がインターフェース2を通じて制御部3に与えられると、制御部3のMPU301は、RAM303に記憶された図8に示したようなディスクアレイ構成情報をその論理アドレスで検索し、仮想ディスク装置番号とディスク内アドレスとを取得する。次に、この取得した仮想ディスク装置番号を持つ仮想ディスク装置の先頭アドレスが、どの実ディスク装置のどの物理アドレスに相当するかを調べるために、前記取得した仮想ディスク装置番号をインターフェース306を通じてアドレス変換部4に送る。

【0032】アドレス変換部4は、受信した仮想ディスク装置番号で図5に示したようなディスク構成テーブルを検索して実ディスク装置番号と通し番号とを取得し、この取得した通し番号とサイズレジスタ5に記憶されている仮想ディスク装置のサイズとから、

(通し番号-1)×サイズ

を計算し、この計算結果と前記取得した実ディスク装置番号とをインターフェース306を通じてMPU301に返却する。

【0033】MPU301は、返却された計算結果に前記ディスク内アドレスを加算することで、返却された実ディスク装置番号を持つ実ディスク装置上でのアドレスを求め、当該実ディスク装置の当該アドレスに対するアクセス処理を行う。まず、キャッシュ302に、当該実ディスク装置の当該アドレスのデータのコピーが存在するか否かを調べ、存在すれば、リード要求であればキャ

(6)

特開2000-298556

9

10

ッシュ302から読み出したデータをインターフェース2を通じてCPU1に返却し、ライト要求であればCPU1からライト要求時に渡されたデータをキャッシュ302に書き込むと同時にセクタ304を介して当該実ディスク装置の当該アドレスにも書き込む。なお、キャッシュ302のみを更新し、実ディスク装置上のデータはキャッシュ302の書き出し時に行うようにしても良い。

【0034】他方、キャッシュ302に、当該実ディスク装置の当該アドレスのデータのコピーが存在しなければ、リード要求の場合は、セクタ304を介して当該実ディスク装置の当該アドレスからデータを読み出してCPU1に返却すると共にキャッシュ302にキャッシングする。ライト要求の場合は、セクタ304を介して当該実ディスク装置の当該アドレスにCPU1からライト要求時に渡されたデータを書き込む。

【0035】(2) RAID1

(A) リード要求

CPU1から論理アドレスを指定したリード要求がインターフェース2を通じて制御部3に与えられると、制御部3のMPU301は、RAM303に記憶された図10に示したようなディスクアレイ構成情報をその論理アドレスで検索し、マスタ側仮想ディスク装置番号とディスク内アドレスとを取得する。次に、この取得したマスタ側仮想ディスク装置番号を持つ仮想ディスク装置の先頭アドレスが、どの実ディスク装置のどの物理アドレスに相当するかをRAID0の場合と同様にアドレス変換部4によって求め、前記マスタ側仮想ディスク装置番号とディスク内アドレスとに対応する、実ディスク装置番号および実ディスク装置内アドレスを求める。そして、当該実ディスク装置の当該アドレスに対するアクセス処理をRAID0の場合と同様に行う。

【0036】(B) ライト要求

CPU1から論理アドレスを指定したライト要求がインターフェース2を通じて制御部3に与えられると、制御部3のMPU301は、RAM303に記憶された図10に示したようなディスクアレイ構成情報をその論理アドレスで検索し、マスタ側仮想ディスク装置番号とスレーブ側仮想ディスク装置番号とディスク内アドレスとを取得する。次に、この取得したマスタ側仮想ディスク装置番号を持つ仮想ディスク装置の先頭アドレスがどの実ディスク装置のどの物理アドレスに相当するか、および、前記取得したスレーブ側仮想ディスク装置番号を持つ仮想ディスク装置の先頭アドレスがどの実ディスク装置のどの物理アドレスに相当するかを、RAID0の場合と同様にアドレス変換部4によって求め、前記マスタ側仮想ディスク装置番号とディスク内アドレスとに対応する実ディスク装置番号および実ディスク装置内アドレス、および、前記スレーブ側仮想ディスク装置番号とディスク内アドレスとに対応する実ディスク装置番号およ

び実ディスク装置内アドレスを求める。そして、当該実ディスク装置の当該アドレスに対する2つのアクセス処理をRAID0の場合と同様に行う。マスタとなる仮想ディスク装置とスレーブとなる仮想ディスク装置とは異なる実ディスク装置上に生成されているので、マスタ、スレーブに対する更新は並行して実施可能である。

【0037】(3) RAID3

(A) リード要求

CPU1から論理アドレスとして、例えばA-1を指定したリード要求がインターフェース2を通じて制御部3に与えられると、制御部3のMPU301は、RAM303に記憶された図12に示したようなディスクアレイ構成情報を論理アドレス(A-x)(xは任意の値)で検索し、論理アドレスA-1~A-6の各々に対応するデータ用仮想ディスク装置番号とディスク内アドレス、および、パリティ用仮想ディスク装置番号とディスク内アドレスとを取得する。

【0038】次に、この取得したデータ用仮想ディスク装置番号、パリティ用仮想ディスク装置番号を持つ仮想ディスク装置の先頭アドレスが、それぞれの実ディスク装置のどの物理アドレスに相当するかをRAID0の場合と同様にアドレス変換部4によって求め、前記データ用仮想ディスク装置番号とディスク内アドレスとに対応する実ディスク装置番号および実ディスク装置内アドレス、前記パリティ用仮想ディスク装置番号とディスク内アドレスとに対応する実ディスク装置番号および実ディスク装置内アドレスを求める。

【0039】そして、当該実ディスク装置の当該アドレスに対するリード処理をRAID0の場合と同様に行って、論理アドレスA-1~A-6に記憶されているデータとそのパリティとを取得する。この際、異なる実ディスク装置上の仮想ディスク装置に対するアクセスは並行して実施することができる。次に、パリティチェックを行い、エラーが無ければ、リード要求された論理アドレスA-1に対応するデータをCPU1に返却し、エラーがあれば、その旨をCPU1に返却する。

【0040】(B) ライト要求

CPU1から論理アドレスとして、例えばA-1を指定したライト要求がインターフェース2を通じて制御部3に与えられると、制御部3のMPU301は、RAM303に記憶された図12に示したようなディスクアレイ構成情報を論理アドレス(A-1)で検索し、論理アドレスA-1に対応するデータ用仮想ディスク装置番号とディスク内アドレス、および、パリティ用仮想ディスク装置番号とディスク内アドレスを取得する。

【0041】次に、この取得したデータ用仮想ディスク装置番号、パリティ用仮想ディスク装置番号を持つ仮想ディスク装置の先頭アドレスが、それぞれの実ディスク装置のどの物理アドレスに相当するかをRAID0の場合と同様にアドレス変換部4によって求め、前記デ

(7)

特開2000-298556

11

12

ータ用仮想ディスク装置番号とディスク内アドレスとに対応する実ディスク装置番号および実ディスク装置内アドレス、前記パリティ用仮想ディスク装置番号とディスク内アドレスとに対応する実ディスク装置番号および実ディスク装置内アドレスを求める。

【0042】そして、当該実ディスク装置の当該アドレスに対するリード処理をRAID0の場合と同様に行って、論理アドレスA-1に記憶されているデータとそのパリティとを取得する。データ用仮想ディスク装置とパリティ用仮想ディスク装置とは異なる実ディスク装置上に生成されているため、並行して処理することが可能である。

【0043】次に、取得した論理アドレスA-1のデータと、ライト要求時にCPU1から渡されたデータと、取得したパリティとから、新たなパリティを生成し、このパリティと、CPU1から渡されたデータとを、RAID0の場合と同様にキャッシュ302または実ディスク装置の元の場所に書き戻す。このときも並行して処理することが可能である。

【0044】図13は本発明のディスクアレイ装置の別の実施例のブロック図である。この実施例が図2に示した実施例と相違するところは、仮想ディスク構成部11内に、CPU1、インターフェース2、制御部3を通じてシステム管理者が任意の値を設定できるサイズ設定レジスタ20を備え、サイズ算出部6は、ディスクアレイに使用する複数の実ディスク装置の容量の最大公約数がサイズ設定レジスタ20に設定された値より小さくなる場合には、サイズ設定レジスタ20に設定された値を仮想ディスク装置のサイズに決定し、小さくない場合は、前記求めた最大公約数を仮想ディスク装置のサイズに決定するようにした点にある。

【0045】例えば、2GBの実ディスク装置と、1.2GBの実ディスク装置との2つの実ディスク装置がある場合、図2の実施例では、その最大公約数である400MBに従って、合計8個の仮想ディスク装置が生成されるが、サイズ設定レジスタ20に例えば1GBを設定しておくと、仮想ディスク装置のサイズは1GBとなり、1GBの仮想ディスクが3つ生成される。これにより、0.2GBは未使用となるが、仮想ディスク装置のサイズが極端に小さくなる事態を避けることができる。

【0046】図14は本発明のディスクアレイ装置の更に別の実施例のブロック図である。この実施例が図2に示した実施例と相違するところは、サイズ算出部6を取り除き、システム管理者がCPU1、インターフェース2および制御部3を通じてサイズレジスタ5に仮想ディスク装置のサイズを直接設定できるようにした点にある。仮想ディスク装置のサイズをシステム管理者自身が決定しなければならないため、システム管理者の作業負担が高まるが、仮想ディスク装置のサイズ、仮想ディスク装置の合計台数をシステム管理者自身が調整できる利

点がある。

【0047】

【発明の効果】以上説明したように、本発明によれば、同じ実ディスク装置上に生成された複数の仮想ディスク装置を含め、同容量の複数の仮想ディスク装置を使ってディスクアレイを構成するため、容量の異なる複数の実ディスク装置を使って全ディスク装置の容量の合計と同じか、それにほぼ近い大容量のディスクアレイを構築することが可能となる。

【図面の簡単な説明】

【図1】本発明の一実施の形態の原理説明図である。

【図2】本発明のディスクアレイ装置の一実施例のブロック図である。

【図3】制御部の一実施例のブロック図である。

【図4】複数の実ディスク装置を使ってディスクアレイを構築する手順の一例を示すフローチャートである。

【図5】ディスク構成テーブルの例を示す図である。

【図6】仮想ディスクと実ディスクとの対応関係を示す図である。

【図7】複数の仮想ディスク装置を使って構成されるRAID0のディスクアレイの構成例を示す図である。

【図8】複数の仮想ディスク装置を使って構成されるRAID0のディスクアレイ用のディスクアレイ構成情報の例を示す図である。

【図9】複数の仮想ディスク装置を使って構成されるRAID1のディスクアレイの構成例を示す図である。

【図10】複数の仮想ディスク装置を使って構成されるRAID1のディスクアレイ用のディスクアレイ構成情報の例を示す図である。

【図11】複数の仮想ディスク装置を使って構成されるRAID3のディスクアレイの構成例を示す図である。

【図12】複数の仮想ディスク装置を使って構成されるRAID3のディスクアレイ用のディスクアレイ構成情報の例を示す図である。

【図13】本発明のディスクアレイ装置の別の実施例のブロック図である。

【図14】本発明のディスクアレイ装置の更に別の実施例のブロック図である。

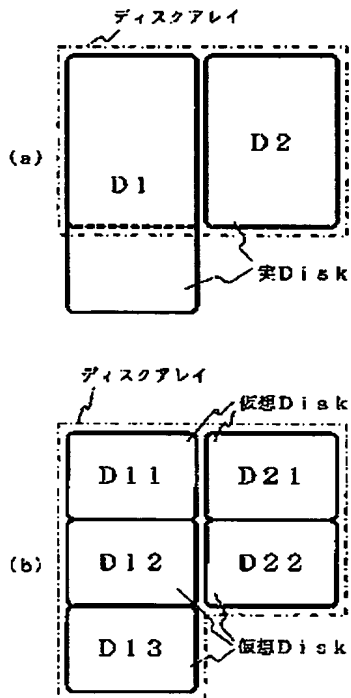
【符号の説明】

- 1…CPU
- 2…インターフェース
- 3…制御部
- 4…アドレス変換部
- 5…サイズレジスタ
- 6…サイズ算出部
- 7…ディスク構成テーブル
- 10…ディスクアレイコントローラ(DAC)
- 11…仮想ディスク構成部
- 20…サイズ設定レジスタ

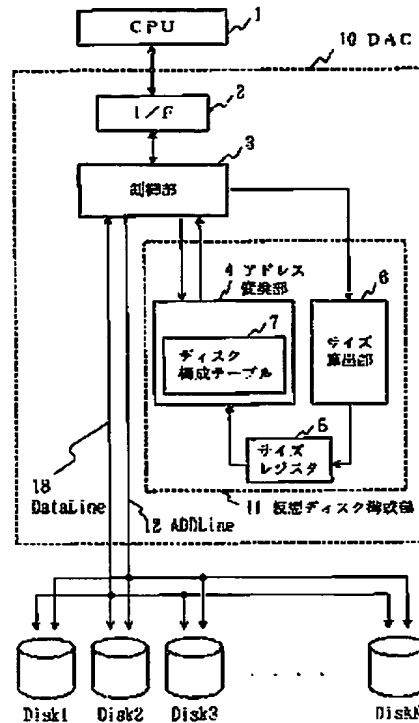
(8)

特開2000-298556

【図1】



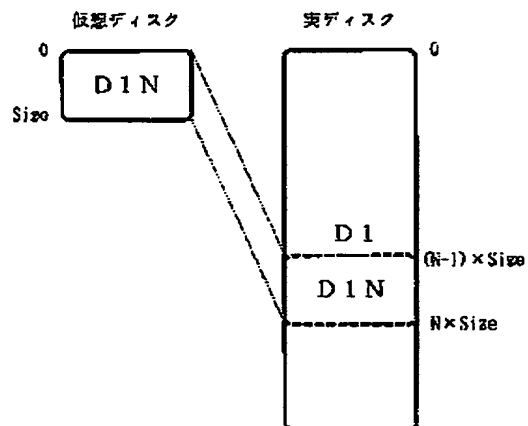
【図2】



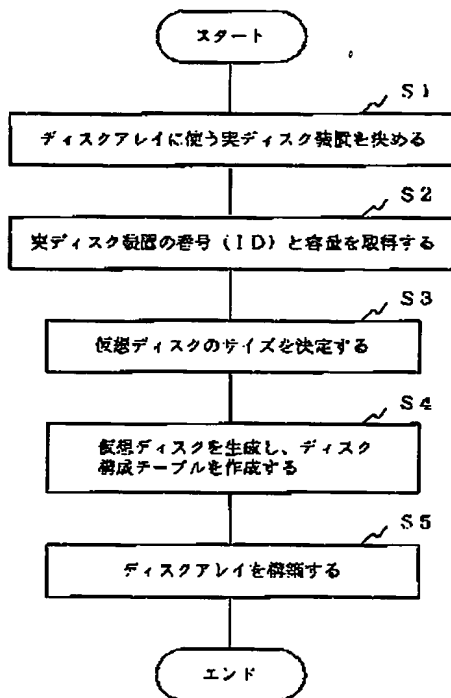
【図5】

仮想Disk	実Disk	No.
D11	D1	1
D12	D1	2
⋮	⋮	⋮
DN1	DN	1
⋮	⋮	⋮
DNM	DN	M

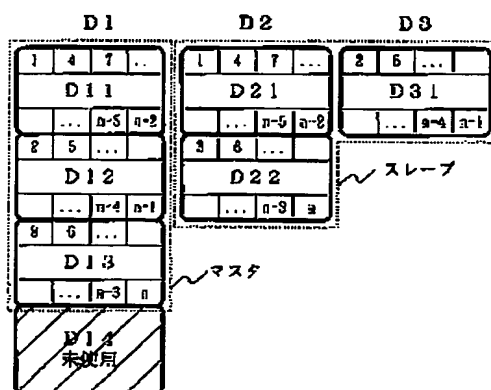
【図6】



【图4】



【圖9】



(10)

特開2000-298556

【図8】

物理アドレス	仮想ディスク装置番号	ディスク内アドレス
1	D11	0
2	D21	
3	D31	
4	D12	
5	D22	
6	D13	
7	D14	
⋮	⋮	⋮
n-6	D11	S-L
n-5	D21	
n-4	D31	
n-3	D12	
n-2	D22	
n-1	D13	
n	D14	

【図10】

物理アドレス	マスター側 仮想ディスク 装置番号	スレーブ側 仮想ディスク 装置番号	ディスク内 アドレス
1	D11	D21	0
2	D12	D31	
3	D13	D22	
4	D11	D21	L
5	D12	D31	
6	D13	D22	
⋮	⋮	⋮	⋮
n-2	D11	D21	S-L
n-1	D12	D31	
n	D13	D22	

【図12】

物理 アドレス	データ用仮想 ディスク装置 番号	データ用仮想 ディスク装置 内アドレス	パリティ用 仮想ディスク 装置番号	パリティ用 仮想ディスク 装置内アドレ ス
A-1	D11	0	D31	G
A-2	D21			
A-3	D12			
A-4	D22	S		L
A-5	D13			
A-6	D14			
B-1	D11	L		
B-2	D21			
B-3	D12			
B-4	D22	S+L		
B-5	D13			
B-6	D14			
⋮	⋮	⋮		⋮
n-1	D11	S-L		S-L
n-2	D21			
n-3	D12			
n-4	D22	2S-L		
n-5	D13			
n-6	D14			

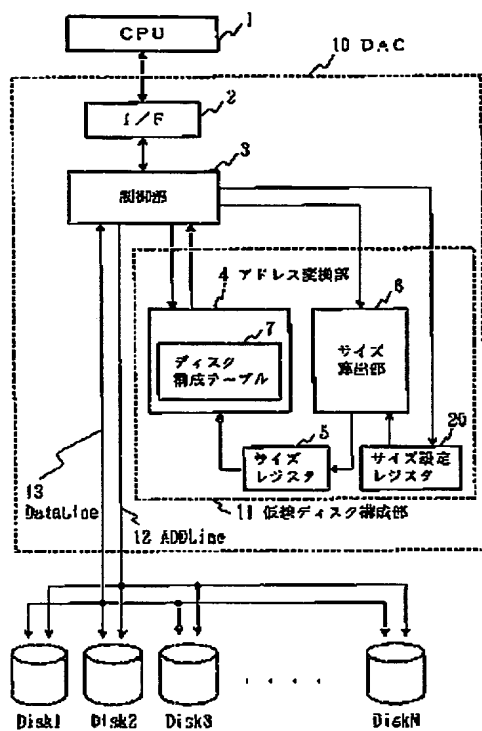
【図11】

D1				D2				D3			
A-1	B-1	C-1	...	A-2	B-2	C-2	...	A-P	B-P	C-P	...
D11				D21				D31			
		...	a-1			...	a-2			...	a-P
A-3	B-3	C-3	...	A-4	B-4	C-4	...				
D12				D22							
		...	a-3			...	a-4				
A-5	B-5	C-5	...								
D13											
		...	a-5								
A-6	B-6	C-6	...								
D14											
		...	a-6								

(11)

特開2000-298556

【図13】



【図14】

